BISS Biodiversity Information Science and Standards

Biodiversity Information Science and Standards 7: e110103
doi: 10.3897/biss.7.110103

OPEN ACCESS

Conference Abstract

# An Overview of the French eDNA Data Landscape: Focus on a national technical repository of reference genetic sequences

Aurélie Lacoeuilhe[‡], Sophie Pamerlon[‡,§], Anne-Sophie Archambeau[‡,§,|], Gael PJ Denys[‡], Yvan Le Bras[‡,¶], Olivier Norvez[‡,¶,#]

‡ Patrinat (OFB-MNHN-CNRS-IRD), 36 rue Geoffroy Saint-Hilaire, Paris, France
§ Global Biodiversity Information Facility (GBIF) France, 43 rue Buffon, Paris, France
| Research Institute for Development (IRD), 44 Bd de Dunkerque, Marseille, France
¶ French virtual research infrastructure in Biodiversity Data (Pôle National de Données de Biodiversité) – PNDB, Station marine de Concarneau Place de la Croix, Concarneau, France
# French Foundation for Biodiversity Research (FRB), Maison des Océans 195 rue Saint-Jacques, Paris, France

Corresponding author: Aurélie Lacoeuilhe (aurelie.lacoeuilhe@mnhn.fr), Sophie Pamerlon (pamerlon@gbif.fr)

Received: 26 Jul 2023 | Published: 27 Jul 2023

Citation: Lacoeuilhe A, Pamerlon S, Archambeau A-S, Denys GP, Le Bras Y, Norvez O (2023) An Overview of the French eDNA Data Landscape: Focus on a national technical repository of reference genetic sequences. Biodiversity Information Science and Standards 7: e110103. https://doi.org/10.3897/biss.7.110103

## Abstract

Environmental DNA (eDNA) data enables biodiversity to be monitored at unprecedented resolution and scale. There is great potential in combining knowledge from traditional and innovative methods such as eDNA for biodiversity assessment. eDNA use cases are increasing in aquatic and marine environments, and studies on soils have been developed in recent years.

PatriNat*[1] (under the guardianship of the French Biodiversity Office (OFB), National Museum of Natural History (MNHN), National Center for Scientific Research (CNRS), and Research Institute for Development (IRD)) is a French data and expertise center working with environmental and research stakeholders to develop data exchanges at all levels. We discuss what eDNA data is (Fig. 1), the different types of data, and the importance of their storage and accessibility.

© Lacoeuilhe A et al. This is an open access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

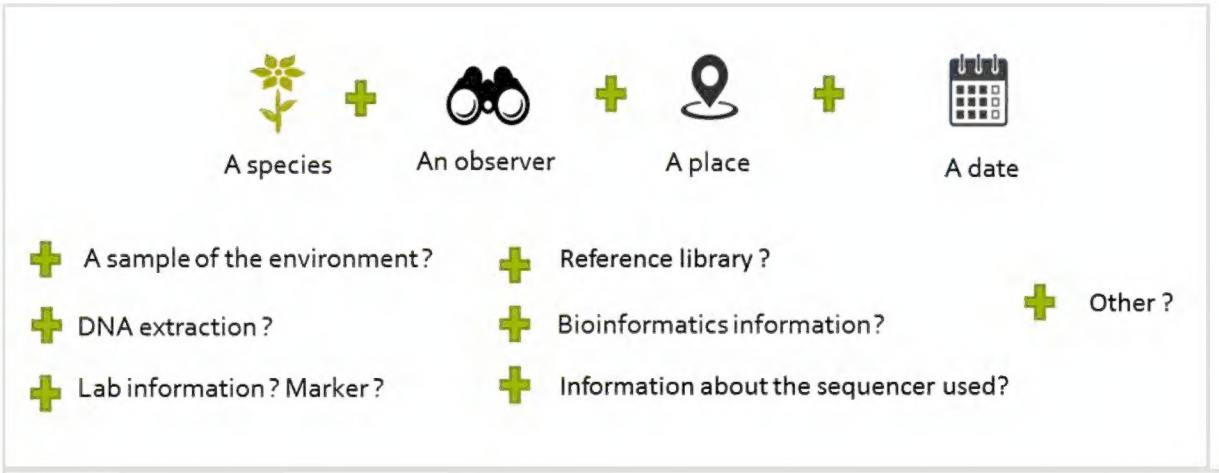


Figure 1.

What is an eDNA exchange data item?

As the amount of eDNA data increases, public agencies need to propose FAIR (Findable, Accessible, Interoperable, Reusable) tools and methods to facilitate their use and foster the development of relevant scientific expertise. We give an overview of the French eDNA data landscape and links with existing standards SINP*[2] (National Heritage Inventory Information System), Darwin Core*[3](Wieczorek et al. 2012) and workflows (Fig. 2).

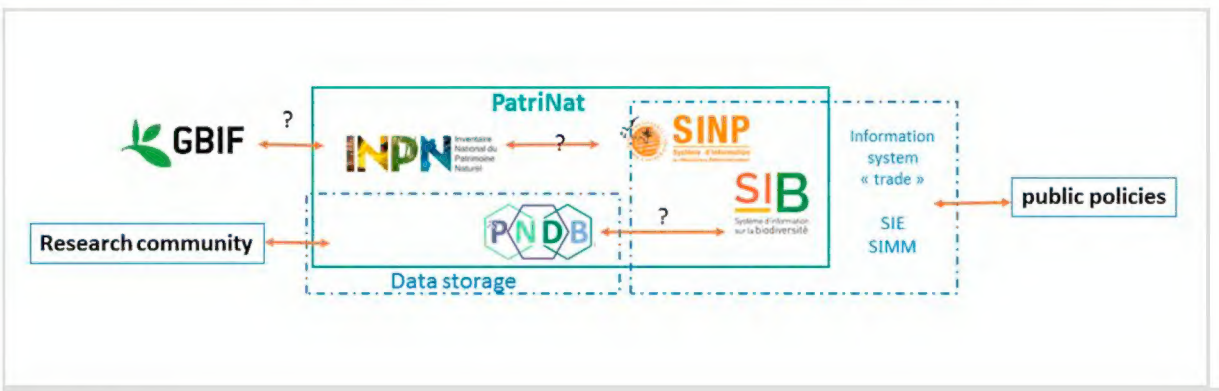


Figure 2.

Links between organizations managing biodiversity and eDNA data. PatriNat hosts the INPN (National Inventory of Natural Heritage), PNDB (French virtual research infrastructure in Biodiversity Data), SIB (Biodiversity information system) and SINP *[2] . Other information systems include SIE (water information system) and SIMM (Marine environment information system).

A priority of eDNA data is to have reliable reference bases and FAIR metadata. PatriNat's new tool will provide access to expertly validated genetic sequence data on species present in France, and is urgently needed for research but also knowledge, monitoring, public policies, and potential law enforcement purposes. We therefore present this technical database built in conjunction with, among other initiatives, DiSSCo*[4], iBOL*[5] and TaxRef*[6] (Gargominy et al. 2022).

It will manage 3 data types:

1. Intrinsic sequence data (marker, sequencing methods, etc.).
2. Sequence management (organization, identifiers, URLs, etc.).
3. Voucher data

It will use the nomenclature of Chakrabarty et al. (2013) as well as:

1. The Global Genome Biodiversity Network (GGBN) standard*[7] (Droege et al. 2016)
2. The Minimum information about any sequence (MIxS) standard (Yilmaz et al. 2011)
3. The Biological Collection Data (ABCD) standard*[8] (Holetschek et al. 2012)
4. The Collections Descriptions terms*[9] (Woodburn et al. 2021)

The DwC DNA extension*[10] can be used for sharing parts of its contents to the Global Biodiversity Information Facility, but referencing sequences associated with specimens/ vouchers will require other standards.

## Keywords

databases, taxonomy, environmental DNA, standards, monitoring biodiversity

## Presenting author

Sophie Pamerlon

## Presented at

TDWG 2023

## Acknowledgements

The authors thank each of the people who took part in discussions on the technical reference system. The work is supported by PatriNat (OFB, MNHN, CNRS, IRD).

## Conflicts of interest

The authors have declared that no competing interests exist.

## References

- Chakrabarty P, Warren M, Page L, Baldwin C (2013) GenSeq: An updated nomenclature and ranking for genetic sequences from type and non-type sources. ZooKeys 346: 29-41. https://doi.org/10.3897/zookeys.346.5753
- Droege G, Barker K, Seberg O, Coddington J, Benson E, Berendsohn WG, Bunk B, Butler C, Cawsey EM, Deck J, Döring M, Flemons P, Gemeinholzer B, Güntsch A, Hollowell T, Kelbert P, Kostadinov I, Kottmann R, Lawlor RT, Lyal C, Mackenzie-Dodds J, Meyer C, Mulcahy D, Nussbeck SY, O'Tuama É, Orrell T, Petersen G, Robertson T, Söhngen C, Whitacre J, Wieczorek J, Yilmaz P, Zetzsche H, Zhang Y, Zhou X (2016)

The Global Genome Biodiversity Network (GGBN) Data Standard specification. Database 2016 https://doi.org/10.1093/database/baw125

- Gargominy O, Tercerie S, Régnier C, Ramage T, Dupont P, Daszkiewicz P, Poncet L (2022) TAXREF, référentiel taxonomique pour la France : méthodologie, mise en œuvre et diffusion. In: Patrinat (OFB - MNHN - CNRS - IRD) (Ed.) Rapport PatriNat (OFB-CNRS-MNHN). Muséum national d'Histoire naturelle, Paris, 47 pp. URL: https://inpn.mnhn.fr/telechargement/referentielEspece/taxref/16.0/menu
- Holetschek J, Dröge G, Güntsch A, Berendsohn WG (2012) The ABCD of primary biodiversity data access. Plant Biosystems - An International Journal Dealing with all Aspects of Plant Biology 146 (4): 771-779. https://doi.org/10.1080/11263504.2012.740085
- Wieczorek J, Bloom D, Guralnick R, Blum S, Döring M, Giovanni R, Robertson T, Vieglais D (2012) Darwin Core: An Evolving Community-Developed Biodiversity Data Standard. PLoS ONE 7 (1). https://doi.org/10.1371/journal.pone.0029715
- Woodburn M, Droege G, Grant S, Groom Q, Jones J, Trekels M, Vincent S, Webbink K (2021) A Data Standard for Dynamic Collection Descriptions. Biodiversity Information Science and Standards 5 https://doi.org/10.3897/biss.5.73902
- Yilmaz P, Kottmann R, Field D, et al. (2011) Minimum information about a marker gene sequence (MIMARKS) and minimum information about any (x) sequence (MIxS) specifications. Nat Biotechnol 29: 415-420. https://doi.org/10.1038/nbt.1823

## Endnotes

*1 https://www.patrinat.fr/fr

*2 https://standards-sinp.mnhn.fr/category/standards/

*3 https://dwc.tdwg.org/

*4 Distributed System of Scientific Collections https://www.dissco.eu/

*5 International Barcode of Life: https://ibol.org/

*6 https://inpn.mnhn.fr/programme/referentiel-taxonomique-taxref

*7 https://www.ggbn.org/ggbn_portal/

*8 https://www.tdwg.org/standards/abcd/

*9 https://www.tdwg.org/community/cd/#resources

*10 https://rs.gbif.org/extension/gbif/1.0/dna_derived_data_2022-02-23.xml